

### Abstract

The Special PRiority and Urgent Computing Environment (SPRUCE) provides on-demand access to high-performance computing resources for time-critical applications. While SPRUCE supports computationally intensive applications, it does not yet fully support the data storage and transfer requirements of these applications. To support data intensive applications in urgent computing environments, we developed the urgent Computing Environment DAta Resource manager (CEDAR). In this poster, we describe the CEDAR architecture and capabilities, illustrate how CEDAR will integrate with urgent computing environments, and evaluate the capabilities of CEDAR in a simulated urgent computing environment.

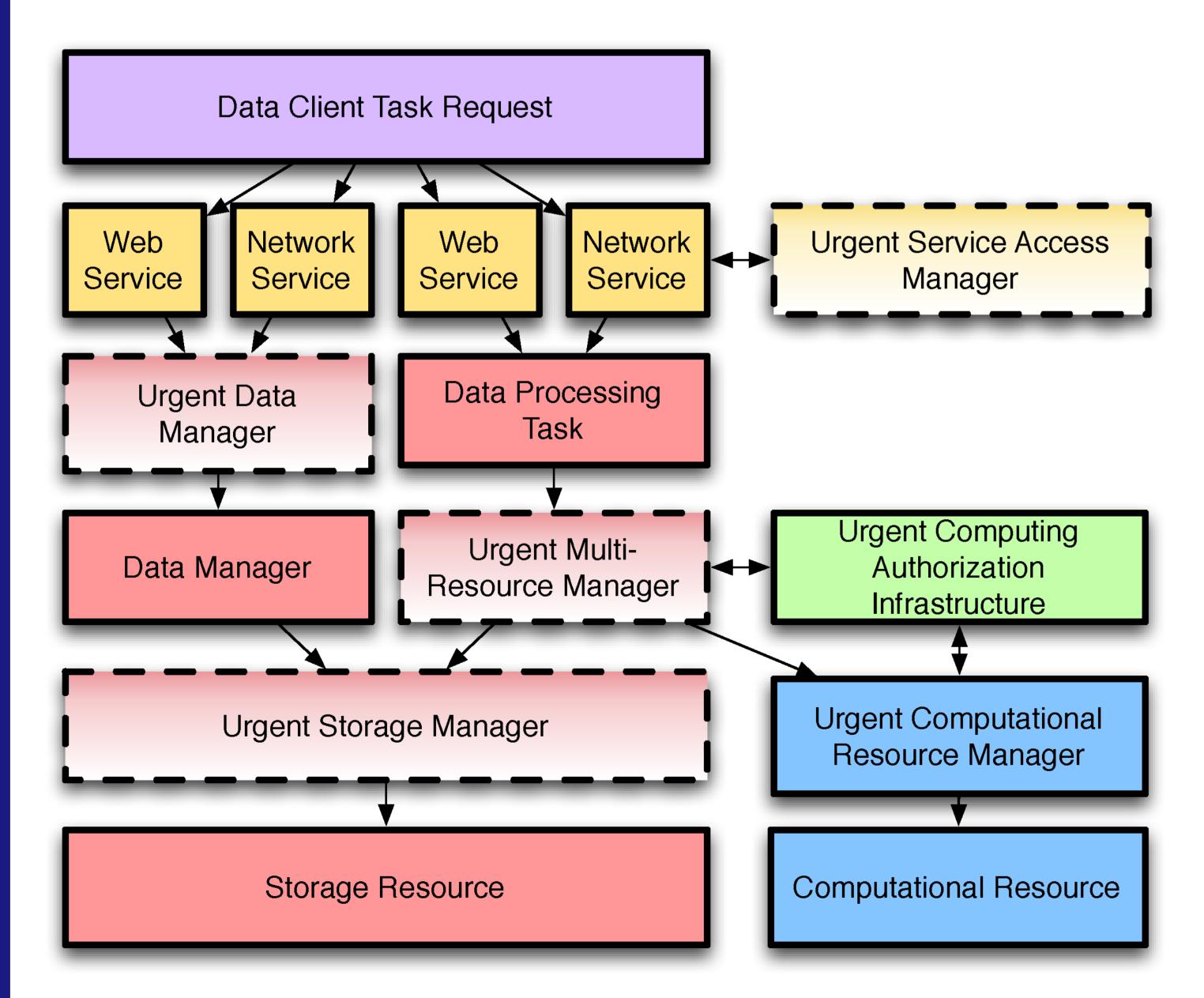
#### Motivation

Urgent computing environments provide cohesive infrastructure to support time-critical, emergency applications executing in shared computing environments, such as the SPRUCE TeraGrid urgent computing environment [1]. Data intensive urgent computing applications require additional resource allocation and provisioning capabilities in urgent computing environments:

- Storage and network resource provisioning tools to ensure the availability or quality of service for a data resource
- Storage and network resource quality of service monitoring and forecasting tools to identify the current and expected state of a data resource
- Automated data management tools that invoke appropriate application or resource data management policies for urgent computing environments

### Urgent Data Management Framework

We are developing the Urgent Data Management Framework (UDMF) to support data intensive urgent computing applications and workflows [2]. The following figure illustrates the UDMF components.



### UDMF provides:

- Data resource provisioning and monitoring capabilities for urgent storage and network resources (red components)
- Urgent Grid service provisioning (yellow components) [3]
- Urgent computing computational and data resource integration (blue components)

# Supporting Storage Resources in Urgent Computing Environments

Jason M. Cope and Henry M. Tufo

Department of Computer Sciences, University of Colorado at Boulder, Boulder, CO, USA {jason.cope, henry.tufo}@colorado.edu

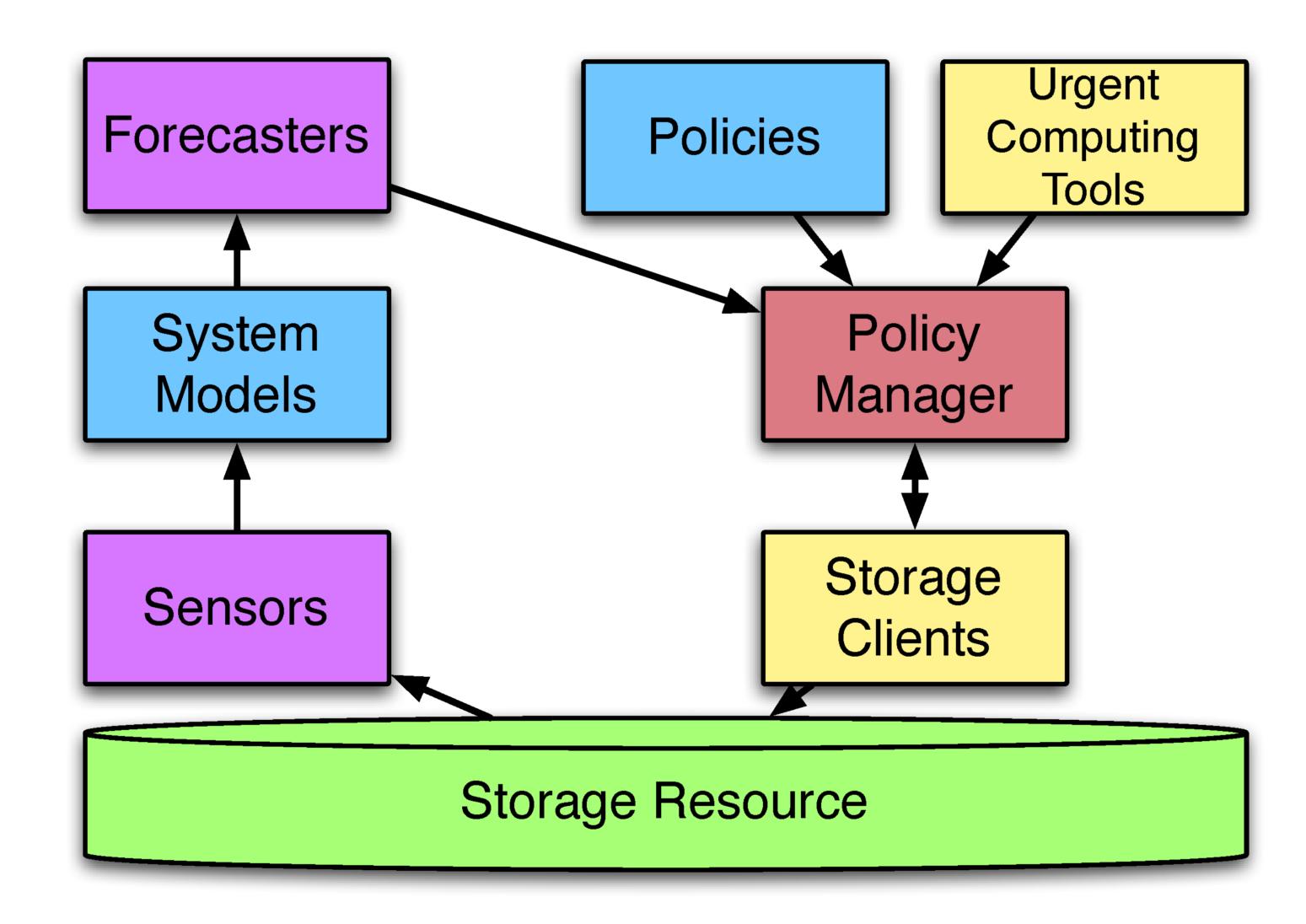
Computer and Information Systems Laboratory, National Center for Atmospheric Research, Boulder, CO, USA {jcope, tufo}@ucar.edu

### CEDAR Architecture and Simulator

The capabilities and features provided by the UDMF urgent Computing Environment DAta Resource manager (CEDAR) include:

- Storage and network provisioning for data resources in urgent computing environments
- Data policy management for urgent computing environments and applications
- Quality of service monitoring for data resources
- Quality of service forecasting for data resources
- Urgent computing environment simulation and analysis

The following figure illustrates the individual components and component interactions within CEDAR.



The storage and network resources are monitored using sensors based on custom resource instrumentation data. This data is used to compute urgent computing environment resource and policy models and forecast the expected state of these models. The forecast data is ingested by a policy manager that pairs a data management policy and data resource with an urgent computing data or application request. The policy manager invokes the selected policy implementation. Policy execution time models currently in the CEDAR policy manager include data compression (1), removal (2), replication (3), and offloading (4) policies.

$$T_C = D(\frac{1}{B_{FS}} + \frac{C_R}{B_{FS}} + \frac{1}{C}) \tag{2}$$

$$T_P = \frac{D_F}{R_R} \tag{2}$$

$$T_R = \frac{D_S N}{min(B_{EG} B_{EG} B_N)} \tag{3}$$

$$T_O = \frac{D_S}{(min(B_{FS_1}, B_{FS_2}, B_N))} + \frac{D_F}{R_R} \tag{4}$$

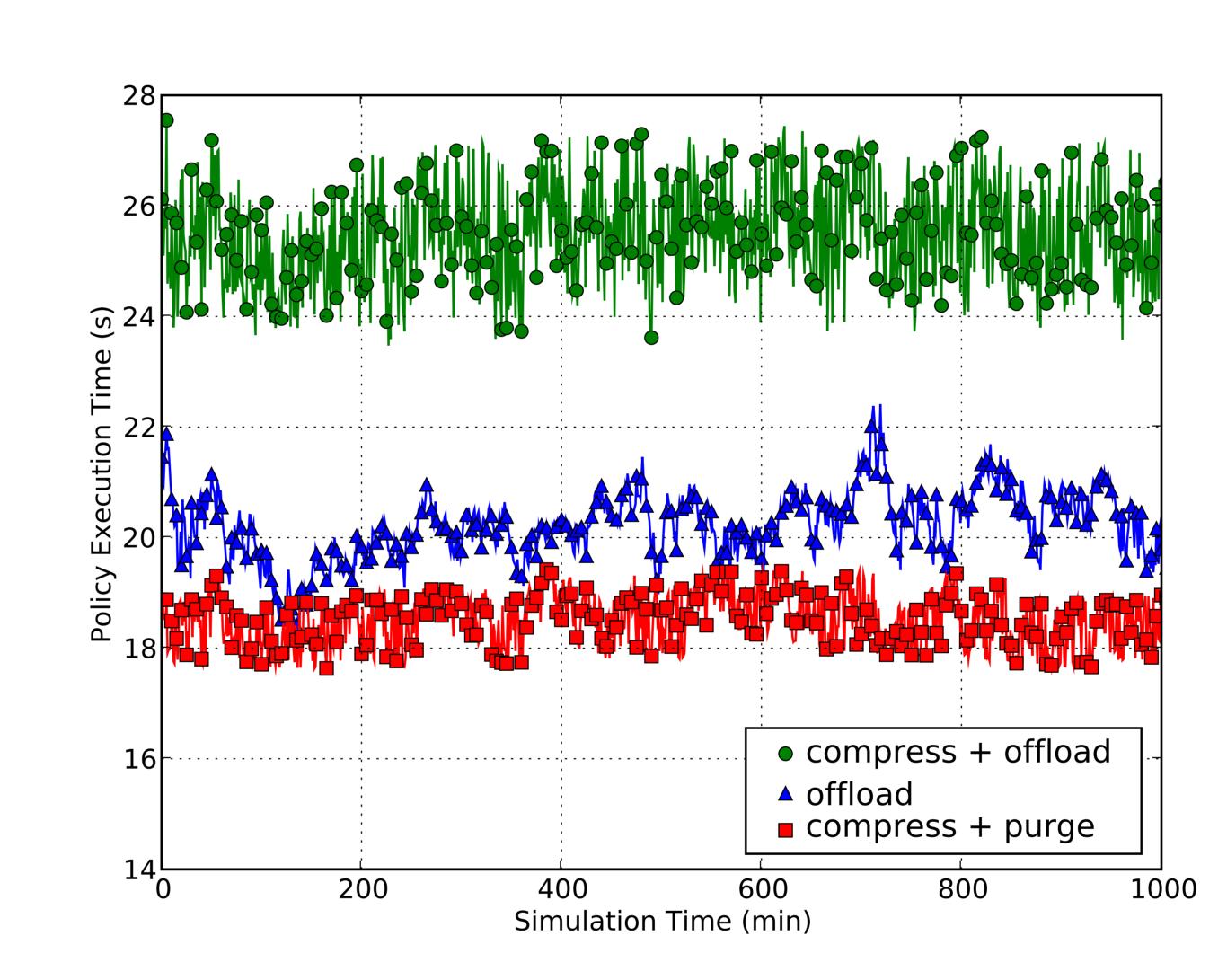
These models are composed of several variables, including the amount of data to be manipulated  $(D_S)$ , the number of files  $(D_F)$ , the number of replicas (N), storage system throughput  $(B_{FS})$ , network throughput  $(B_N)$ , data compression ratio  $(C_R)$  and rates (C), and data removal rates  $(R_R)$ . The result of the models is the time required to complete the policy (T).

### CEDAR Data Management Policy Simulation

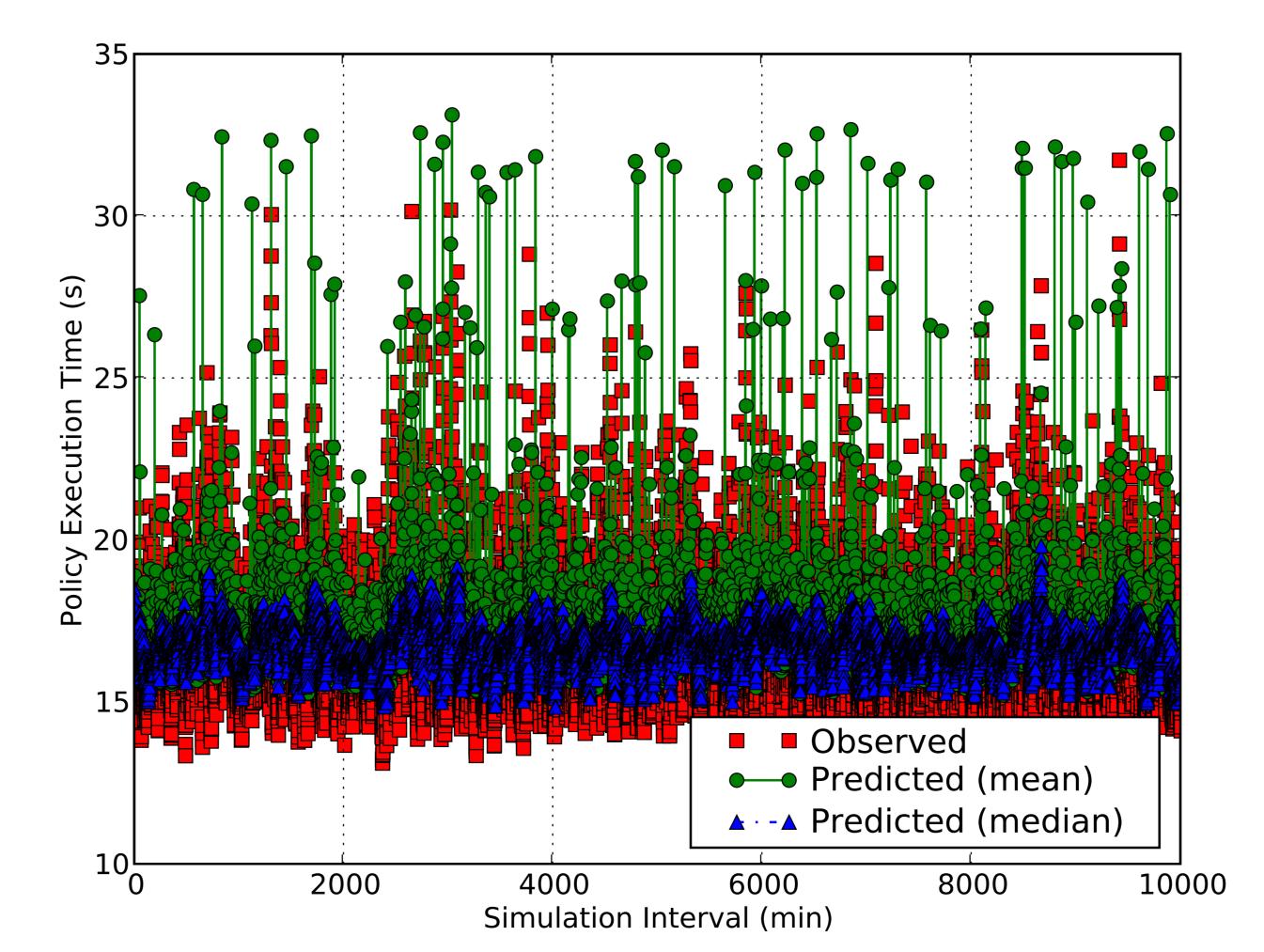
Goals of the data management policy simulation:

- Demonstrate the CEDAR policy and urgent computing environment simulation and analysis capabilities
- Evaluate several simple CEDAR data management policies
- Evaluate several simple data management policy selection tools for CEDAR

For this evaluation, we setup the CEDAR simulator in random event generation mode using storage availability, storage bandwidth, and network bandwidth data distributions generated from data collected from the NCAR Frost, NCSA Mercury, and TACC Lonestar TeraGrid resources between April 2008 and July 2008 and static values ( $R_R = 300$ ,  $C_R = 0.5$ , and  $D_S = 1$ GB). The following figure illustrates the amount of time required to execute offload, compress+removal, and compress+offload policies.



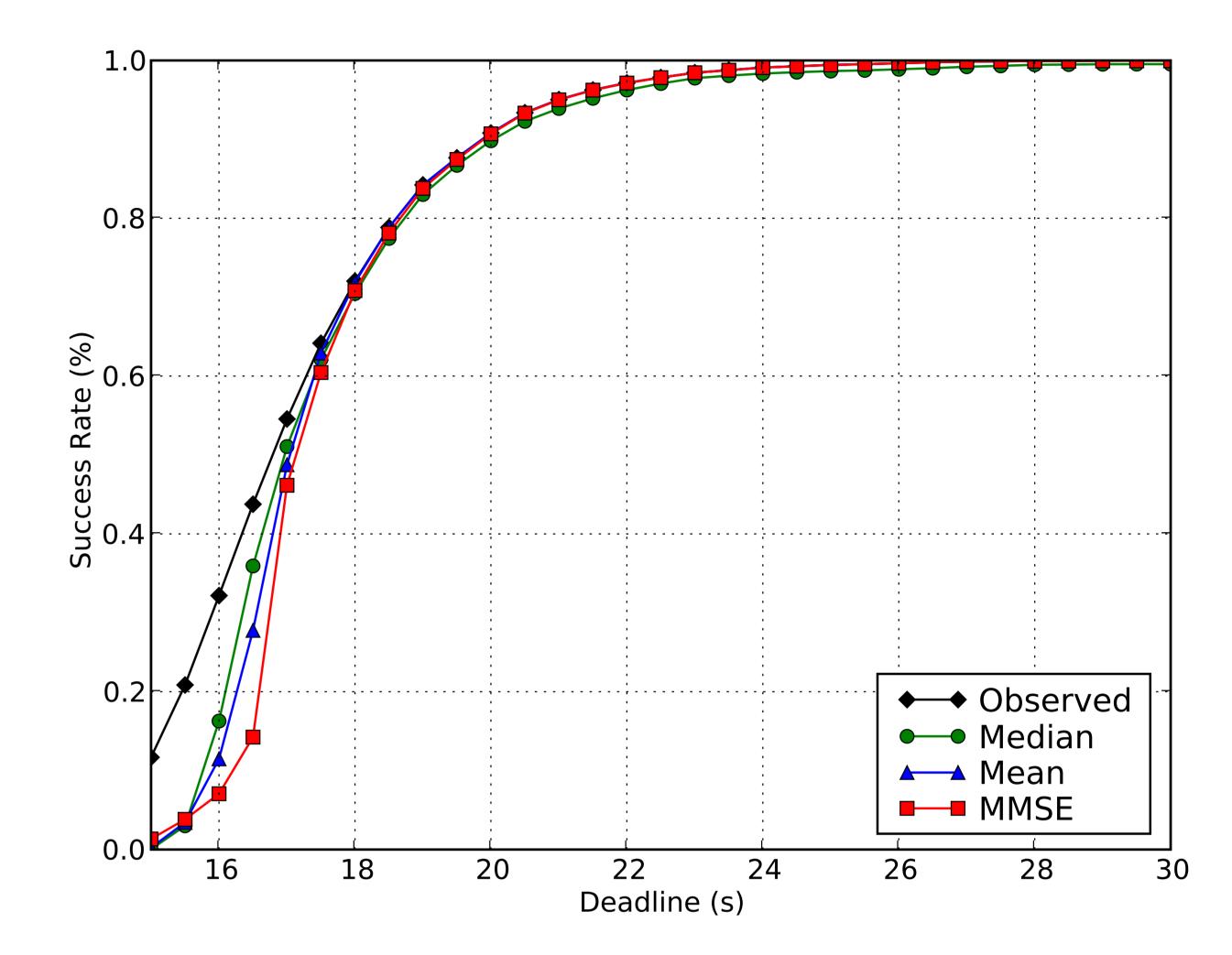
The following figure illustrates the variable performance of mean and median based policy prediction methods overlaid on the observed performance of the data replication policy. More useful forecasts could be derived from upper or lower bound parameter predictions.





## NCAR

Our simulation results evaluating how often each predictor provided a correct policy forecast that satisfied the urgent computing deadline requirement are below.



The mean, median, and minimum mean square error (MSSE) predictors were used in this evaluation. Depending upon application deadline requirements, specific policy predictors may be more accurate than others. From this evaluation, we observed that:

- The median-based method outperforms the MMSE selection method when 15s 19s deadlines are required
- The MMSE method outperforms the median-based method from the 19s 30s deadline interval range

We made similar observations on the performance of policy predictors based on individual urgent computing resource parameters (such as the  $B_N$  policy parameter). These parameters can influence policy selection and will be accounted for in the policy selection tools.

### Conclusions and Future Work

To better prepare and adapt data resources for urgent computing environments, we have developed a data management system and simulator to evaluate urgent computing data management policy, application, and resource interactions. We have identified several areas of improvement and future work:

- Predicting the upper or lower bounds of policy execution may better identify the expected execution limits of a policy
- Additional policy and resource selection tools are required to identify the best policies and resources that are tolerant to changes in policy parameters and account for changes in policy or environment behavior
- [1] P. Beckman, I. Beschatnikh, S. Nadella, and N. Trebon. Building an Infrastructure for Urgent Computing. *High Performance Computing and Grids in Action*, 2007.
- [2] J. Cope and H. Tufo. A Data Management Framework for Urgent Geoscience Workflows. In *Proceedings of the International Conference on Computational Science (ICCS 2008)*, June 2008.
- [3] J. Cope and H. Tufo. Adapting Grid Services for Urgent Computing Environments. In *Proceedings of the 2008 International Conference on Software and Data Technologies (ICSOFT 2008)*, July 2008.

Contact information: Jason M. Cope, 430 UCB, Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309 – Email: jason.cope@colorado.edu; Web: http://www.csc.cs.colorado.edu/copej/ Acknowledgments: University of Colorado computer time was provided by equipment supported under National Center for Atmospheric Research, NSF MRI Grant #CNS-0420873, NSF MRI Grant #CNS-0420873, NSF MRI Grant #CNS-0420873, NSF MRI Grant #CNS-0420985, NSF MRI Grant #CNS-0420873, NSF MRI Grant #CNS-04